Seminar on Resampling Methods

Carsten Jentsch

Business and Social Statistics
Department of Statistics
TU Dortmund University

Winter Term 2024/25

Time Schedule

Seminar kick-off meeting:

- **First** week of the (upcoming) semester (exact time and date will be announced later; in agreement with participants)

Meeting to assign seminar topics to participants:

- **Second** week of the semester (exact time and date will be announced later; in agreement with participants)

Seminar presentations:

- Block seminar **at the end** of (after) the semester (exact time and date in agreement with participants)

What are resampling methods?

Resampling (Statistics) (engl. Wikipedia):

In statistics, **resampling** is any of a variety of methods for doing one of the following:

1.) Estimating the precision of sample statistics (medians, variances, percentiles) by using subsets of available data (**jackknifing, subsampling**) or drawing randomly with replacement from a set of data points (**bootstrapping**)

2.) Exchanging labels on data points when performing significance tests (**permutation tests**, **randomization tests**, or **re-randomization tests**)

3.) Validating models by using random subsets (**bootstrapping**, **cross validation**)

Universal technique(s) to approach various tasks:

- statistical inference (in general), confidence intervals, testing, model selection, variable selection, tuning parameter selection, prediction, etc.

Possible topics

Modern statistical methodology using resampling methods in different contexts:

- resampling for big data
- resampling for causal inference
- multiplier / dependent wild bootstrapping
- subsampling
- frequency domain bootstrapping
- resampling / bootstrapping for prediction
- bootstrapping empirical processes
- bootstrapping with (lasso) penalization
- bootstrapping for functional data
- Bayesian bootstrap
- ...

Possible (concrete) topics

Causality:

- Imbens, G. and Menzel, K. (2021). A causal bootstrap. Ann. Statist. 49 (3), 1460 - 1488.
- Little, M. and Badawy, R. Causal bootstrapping. arXiv:1910.09648v3
- Lia, J. (2023). Are we bootstrapping the right thing? A new approach to quantify uncertainty of Average Treatment Effect Estimate. arXiv:2310.11683v4+

Spatial:

- Rabinowiczy, A. and Rosset, S. (2021) Resampling Methods for Detecting Anisotropic Correlation Structure
- Arthur, R. (2024). A General Method for Resampling Autocorrelated Spatial Data. arXiv:2401.05728v1

Possible (concrete) topics

Privacy:

- Pelaeza, R., Aneirosa, G. and Vilara, J. (2024). Bootstrap prediction regions for daily curves of electricity demand and price using functional data. arXiv:2401.11885v1
- Dette, H. and Graw, C. (2024). Uncertainty quantification by block bootstrap for differentially private stochastic gradient descent. arXiv:2405.12553v1
- Chadha, K., Duchi, J. and Kuditipudi, R. (2024). Resampling methods for private statistical inference. arXiv:2402.07131v2

Possible (concrete) topics

Time Series:

- La Vecchia, D., Moor, A. and Scaillet, O. (2022). A Higher-Order Correct Fast Moving-Average Bootstrap for Dependent Data. arXiv:2001.04867v2
- Betken, A., Dehling, H. and Kroll, M. (2021). Block Bootstrapping the empirical distance covariance. arXiv:2112.14091v1
- Cho, H. and Kirch, C. (2021). Bootstrap confidence intervals for multiple change points based on moving sum procedures. arXiv:2106.12844v1
- Sirotko-Sibirskaya, N., Franz, M. O. and Dickhaus, T. (2020). Volterra bootstrap: Resampling higher-order statistics for strictly stationary univariate time series. arXiv:2010.10071v1
- Franke, J. and Nyarige, E. G. (2019). A residual-based bootstrap for functional autoregressions. arXiv:1905.07635v1
- Palm, N. Nagler, T. (2023). An online bootstrap for time series. arXiv:2310.19683v

Possible (concrete) topics

Methods:

- Cavaliere, G. and Georgiev, I. (2020). Inference under random limit Bootstrap measures. Econometrica 88 (2), 2547 - 2574.
- Hong, H. and Li, J. (2020). The numerical Bootstrap. Ann. Statist. 48 (1), 397 - 412.
- Cavaliere, G., Goncalves, S. (2022). Bootstrap Inference in the Presence of Bias. arXiv:2208.02028v1
- Giessing, A. and Fan, J. (2023). A Bootstrap Hypothesis Test for High-Dimensional Mean Vectors arXiv:2309.01254v1
- Robbins, M. Burgette, L. and Bauhoff, S. (2023). Resampling Methods with Imputed Data. arXiv:2311.13815v1
- Zrimsek, U. and Strumbelj, E. (2024). Quantifying Uncertainty: All We Need is the Bootstrap? arXiv:2403.20182v1

Possible (concrete) topics

Computation:

- Lunde, R., Sarkar, P. and Ward, R. (2021). Bootstrapping the error of Oja's algorithm. arXiv:2106.14857v1
- Ramprasad, P., Li, Y., Yang, Z. and Wang, Z. (2021). Online Bootstrap Inference For Policy Evaluation In Reinforcement Learning. arXiv:2108.03706v3
- Ma, Y. Leng, C. and Wang, H. (2023). Optimal Subsampling Bootstrap for Massive Data. arXiv:2302.07533v1
- Liu, K., Blanchet, J., Ying, L. and Lu, Y. (2024). Orthogonal Bootstrap: Efficient Simulation of Input Uncertainty. arXiv:2404.19145v1

Introductory reading

- Chernick (2008). Bootstrap Methods - Guide for Practitioners and Researchers
- Davison & Hinkley (1997). Bootstrap Methods and Their Applications
- Dikta & Scheer (2021). Bootstrap Methods: With Applications in R.
- Efron & Tibshirani (1993). An Introduction to Bootstrap.
- Lahiri (2003). Resampling Methods for dependent data.
- Politis, Romano & Wolf (1999). Subsampling.
- Shao & Tu (1993). The Jackknife and Bootstrap.

Requirements to pass the seminar

Bachelor:

- presentation (30 minutes)
- seminar paper (10 pages)
- active participation in discussions, feedback

Master:

- presentation (45 minutes)
- seminar paper (10 pages)
- active participation in discussions, feedback

For question, please write an email to

fioravanti@statistik.tu-dortmund.de
or
jentsch@statistik.tu-dortmund.de